METHOD FOR TWO PARTY AUTHENTICATION AND KEY AGREEMENT

RELATED APPLICATIONS

The following applications, filed concurrently with the subject application, are related to the subject application and are hereby incorporated by reference in their entirety: application no. unknown entitled METHOD FOR UPDATING SECRET SHARED DATA IN A WIRELESS COMMUNICATION SYSTEM by the inventor of the subject application; application no. unknown entitled METHOD FOR TRANSFERRING SENSITIVE INFORMATION USING INTIALLY UNSECURED COMMUNICATION by the inventor of the subject application; application no. unknown entitled METHOD FOR SECURING OVER-THE-AIR COMMUNICATION IN A WIRELESS SYSTEM by the inventor of the subject application; and application no. unknown entitled METHOD FOR ESTABLISHING A KEY USING OVER-THE-AIR COMMUNICATION AND PASSWORD PROTOCOL AND PASSWORD PROTOCOL by the inventor of the subject application and Adam Berenzweig.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a method for authenticating parties communicating with one another, and in one application, a method for authenticating a mobile and a network in wireless communication. The present invention further relates to a key agreement based on the authentication protocol.

2. Description of Related Art

Protocols for authenticating parties communicating with one another provide a measure of security to the communication. Several such protocols are employed by the wireless industry and form part of the different communication standards in the U.S., Europe and Japan.

While the party authentication system and method according to the present invention is not limited to wireless communication, to promote ease of understanding, the present invention will described in the context of a wireless system. For this reason, a general overview of wireless systems is presented, including a discussion of the party authentication protocol used in at least one of the standards.

The U.S. currently utilizes three major wireless systems, with differing standards. The first system is a time division multiple access system (TDMA) and is governed by IS-136, the second system is a code division multiple access (CDMA) system governed by IS-95, and the third is the Advanced Mobile Phone System (AMPS). All three communication systems use the IS-41 standard for intersystem messaging, which defines the authentication

procedure for call origination, updating the secret shared data, and etc.

Fig.1 illustrates a wireless system including an authentication center (AC) and a home location register (HLR) 10, a visiting location register (VLR) 15, and a mobile 20. While more than one HLR may be associated with an AC, currently a one-to-one correspondence exists. Consequently, Fig. 1 illustrates the HLR and AC as a single entity, even though they are separate. Furthermore, for simplicity, the remainder of the specification will refer to the HLR and AC jointly as the AC/HLR. Also, the VLR sends information to one of a plurality of mobile switching centers (MSCs) associated therewith, and each MSC sends the information to one of a plurality of base stations (BSs) for transmission to the mobile. For simplicity, the VLR, MSCs and BSs will be referred to and illustrated as a VLR. Collectively, the ACs, HLRs, VLRs, MSCs, and BSs operated by a network provider are referred to as a network.

A root key, known as the A-key, is stored only in the AC/HLR 10 and the mobile 20. There is a secondary key, known as Shared Secret Data SSD, which is sent to the VLR 15 as the mobile roams (i.e., when the mobile is outside its home coverage area). SSD is generated from the A-key and a random seed RANDSSD using a cryptographic algorithm or function. A cryptographic function is a function which generates an output having a predetermined number of bits based on a range of possible inputs. A keyed cryptographic function (KCF) is a type of cryptographic function that operates based on a key; for instance, a cryptographic function which operates on two or more arguments (i.e., inputs) wherein one of the arguments is

the key. From the output and knowledge of the KCF in use, the inputs can not be determined unless the key is known. Encryption/decryption algorithms are types of cryptographic functions. So are one-way functions like pseudo random functions (PRFs) and message authentication codes (MACs). The expression $KCF_{SK}(R_N')$ represents the KCF of the random number R_N' using the session key SK as the key. A session key is a key that lasts for a session, and a session is a period of time such as the length of a call.

In the IS-41 protocol, the cryptographic function used is CAVE (Cellular Authentication and Voice Encryption). When the mobile 20 roams, the VLR 15 in that area sends an authentication request to the AC/HLR 10, which responds by sending that mobile's SSD. Once the VLR 15 has the SSD, it can authenticate the mobile 20 independently of the AC/HLR 10. For security reasons, the SSD is periodically updated.

Fig. 2 illustrates the communication between the AC/HLR 10, the VLR 15 and the mobile 20 to update the SSD. As discussed above, the AC/HLR 10 generates a random number seed RANDSSD, and using the CAVE algorithm generates a new SSD using the random number seed RANDSSD. The SSD is 128 bits long. The first 64 bits serve as a first SSD, referred to as SSDA, and the second 64 bits serve as a second SSD, referred to as SSDB. As shown in Fig. 2, the AC/HLR 10 provides the VLR 15 with the new SSD and the RANDSSD. The VLR 15 then sends the RANDSSD to the mobile 20 along with a session request SR. The session request SR instructs the mobile 20 to perform the SSD update protocol which is described in detail below. In response to receipt of the RANDSSD and the session request

SR, the mobile 20 uses the CAVE algorithm to generate the new SSD using the RANDSSD, and generates a random number R_M using a random number generator. The mobile sends the random number R_M to the VLR 15. The mobile 20 also performs the CAVE algorithm on the random number R_M using the new SSDA as the key. This calculation is represented by $CAVE_{SSDA}(R_M)$.

One of the VLR 15 and the AC/HLR 10, also calculates $CAVE_{SSDA}(R_M)$, and sends the result to the mobile 20. The mobile 20 authenticates the network if $CAVE_{SSDA}(R_M)$, which it calculated, matches that received from the network.

Next, and usually after receiving a signal from the mobile 20 indicating verification, the VLR 15 generates a random number R_N, and sends the random number R_N to the mobile 20. Meanwhile, the VLR calculates $CAVE_{SSDA}(R_N)$. Upon receipt of R_N, the mobile 20 calculates $CAVE_{SSDA}(R_N)$, and sends the result to the VLR 15. The VLR 15 authenticates the mobile if $CAVE_{SSDA}(R_N)$, which it calculated, matches that received from the mobile 20. The random numbers R_M and R_N are referred to as challenges, while $CAVE_{SSDA}(R_M)$ and $CAVE_{SSDA}(R_N)$ are referred to as challenge responses. Once the authentication is complete, the mobile 20 and the network generate session keys using SSDB.

In this protocol, the SSD is itself used to answer the challenges from the mobile 20 and the network. This allows an attack when an old RANDSSD and SSD pair are revealed. Knowing this pair is enough to query the mobile 20, and answer its challenge. Thus an attacker can issue an SSD update to the mobile 20, and answer the challenge from the mobile. Once the revealed SSD is accepted, and despite a secure session key agreement protocol (i.e., a protocol on communication between a

mobile and a network to establish a session key), the attacker can impersonate the network and place a call to the mobile 20 under fraudulent identities. For example, the impersonator can insert his own caller id or name and pretend to be someone else. The attacker can pretend to be a credit card company, and ask to verify card number and pin. Or even use the telephone company name in the caller name field and ask to verify calling card numbers, etc.

SUMMARY OF THE INVENTION

In a two party authentication method according to the present invention a first party issues a random number as a first challenge, and a second party responds with a first challenge response. The first challenge response is generated by performing a keyed cryptographic function (KCF) on the first challenge and a count value using a first key. The second party increments the count value upon receipt of the first challenge, and uses the count value as a second challenge. The first party verifies the second party based on the first challenge and receipt of the second challenge and the first challenge response. After verification, the first party performs the KCF on the second challenge using the first key to generate a second challenge response. Based on the second challenge and receipt of the second challenge response, the second party verifies the first party. Using the first and second challenges, an encryption key is generated by both parties. In this manner, a different key, the first key, from the encryption key is used in answering challenges. The present invention has many applications including the wireless industry wherein the

first and second parties are a network and mobile, respectively.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will become more fully understood from the detailed description given below and the accompanying drawings which are given by way of illustration only, wherein like reference numerals designate corresponding parts in the various drawings, and wherein:

Fig. 1 is a block diagram illustrating the basic components of a wireless system;

Fig. 2 illustrates the communication between the authentication center/home location register, visiting location register, and the mobile to update the secret shared data according to the IS-41 standard; and

Fig. 3 illustrates the communication between the network and the mobile to authenticate these two parties according to one embodiment of the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

As discussed above, while the party authentication system and method according to the present invention is not limited to wireless communication, to promote ease of understanding, the present invention will described in the context of a wireless system. More specifically, the method or protocol for two party authentication according to the present invention will be described as employed by the wireless system shown in Fig. 1.

In contrast to the method or protocol discussed above with respect to Figs. 1 and 2, in the method according to the present invention, the AC/HLR 10 and the

mobile 20 also generate another key, referred to as the M-key, based on the A-key. For example, the M-key is generated by applying a pseudo random function (PRF) indexed by the A-key on a value known to the network and the mobile 20. A practical PRF is the well-known Data Encryption Standard-Cipher Block Chaining (DES-CBC) algorithm from NIST (National Institute of Standards). In a preferred embodiment, DES-CBC, indexed by the 64-bit A-key on a known value, produces a 64-bit M-key.

Fig. 3 illustrates the communication between the network and the mobile 20 to authenticate these two parties according to one embodiment of the present invention. As shown, the VLR 15 acts as a conduit for communication between the AC/HLR 10 and the mobile 20. More specifically, the authentication protocol according to the present invention is performed between the AC and the mobile 20.

As shown, one party, the AC/HLR 10, generates and sends a random number R_N to the other party, the mobile 20. Typically, the AC/HLR 10, in addition to sending the random number R_N, sends a session request SR specifying the type of protocol to be performed. Protocol types include, for example, call origination, secret shared data (SSD) update, call termination, and mobile registration.

In response, the mobile 20 generates a count value C_M, and performs a KCF on the random number R_N, the count value C_M, Type data, and id data ID_M using the M-key as the key. This calculation is represented a $KCF_{M\text{-}Key}(Type, ID_M, C_M, R_N)$. Preferably, the KCF is a keyed message authentication code such as HMAC, but could be a PRF such as DES-CBC. The mobile 20 includes a counter which generates the count value C_M. The mobile 20 increments the count value C_M prior

to generating the challenge response (i.e., KCF_{M-Key}(Type, ID_M, C_M , R_N)) to each challenge from the network.

The Type data represents the type of protocol being performed. The id data indicates that the communication issued from the mobile. Usually, id data 1 indicates that the communication is from the network, and id data 0 indicates that the communication came from the mobile. For the purposes of discussion, however, the id data for the mobile 20 is shown as ID_M and the id data for the network is shown as ID_N. The system and method for two party authentication does not require the inclusion of the Type data when performing the KCF on the random number R_N and the count value C_M. The Type data and the specific id data have been included as part of the application of the two party authentication method and system to a wireless system.

The mobile 20 sends count value C_M and KCF_{M-Key}(Type, ID_M, C_M , R_N) to the network. Because the AC/HLR 10 initiated the current protocol including the two party authentication protocol according to the present invention, the AC/HLR 10 knows the Type data. Also, because communication from mobiles include the same id data, this value is known by the AC/HLR 10 as well. Accordingly, based on the received count value C_M, the AC/HLR 10 calculates KCF_{M-Key}(Type, ID_M, C_M , R_N) and determines whether this calculated value matches the version received from the mobile 20. If a match is found, the AC/HLR 10 authenticates the mobile 20.

Once the mobile 20 has been verified, the AC/HLR 10 calculates KCF_{M-Key}(Type, ID_N, C_M), and sends the calculated result to the mobile 20. The mobile 20, meanwhile, calculates KCF_{M-Key}(Type, ID_N, C_M) as well. The mobile 20

then verifies whether the calculated version of $KCF_{M\text{-}Key}$(Type, ID_N C_M) matches the version received from the AC/HLR 10. If a match is found, the mobile 20 authenticates the network.

Furthermore, after performing this two party authentication protocol, other keys can be generated. For instance, if the wireless system of Fig. 1 used this two party authentication protocol as part of the SSD update protocol, then, after the mobile 20 authenticates the network, the mobile 20 and the AC/HLR 10 both have the random number R_N and the count value C_M . Both the mobile 20 and AC/HLR 10 generate the SSD as $PRF_{A\text{-}Key}(C_M\ ,R_N)$; wherein the PRF is preferably the DES-CBC algorithm. Alternatively, in other protocols, this same technique is used to generate other keys.

When applied to a wireless system, the mobile 20 needs to store the count value C_M in semi-permanent memory so that during power down, the count value C_M is not re-initialized. This way, repetition of a count value is prevented; repetition of the count value permits an attacker to prevail in his attack. In a preferred embodiment, the count value is initialized using a random number and generated using a large bit counter such as a 64 or 75 bit counter. This provides security even when the mobile 20 crashes and loses the stored count value. Even if an attacker can cause a mobile to crash at will, and assuming it takes at least a second to initiate a session, it will take, for example, a year before the attacker manages to have the mobile repeat a count value when a 75 bit counter is used.

As a further alternative, instead of a sending a unique random number R_N, the initiating party (e.g., the

network) issues a global random number. Namely, for each communication, the initiating party issues a different, unique random number R_N in the embodiment of Fig. 3. However, in this alternative embodiment, the initiating party issues the same random number R_N for each communication.

In the protocol according to the present invention the key previously established between the parties (e.g., A-key or SSD) is not used to answer challenges, and thus the network impersonation problem discussed with respect to IS41 is not possible. Furthermore, even if the M-key is revealed to an attacker, there is no direct way to obtain the A-key therefrom because a one-way function was used to generate the M-key. Because an attacker uses prior challenges and challenge responses when mounting an attack, such an attack will fail if made on the protocol according to the present invention. The reason is that the attacker will be using a challenge response based on an old count value. Consequently, the network will not verify the attacker. Further, keys generated after authentication as discussed above will be generated by the PRF of the new challenge using the A-key, and the attacker does not know the A-key.

The invention being thus described, it will be obvious that the same may be varied in many ways. Such variations are not to be regarded as a departure from the spirit and scope of the invention, and all such modifications are intended to be included within the scope of the following claims.